

Intrinsic noise of microRNA-regulated genes and the ceRNA hypothesis

Javad Noorbakhsh¹, Alex Lang¹, Pankaj Mehta^{1,*}

¹ Physics Department, Boston University, Boston, MA, USA

* E-mail: pankajm@bu.edu

MicroRNAs are small noncoding RNAs that regulate genes post-transcriptionally by binding and degrading target eukaryotic mRNAs. We use a quantitative model to study gene regulation by microRNAs and compare it to gene regulation by prokaryotic small non-coding RNAs (sRNAs). Our model uses a combination of analytic techniques as well as computational simulations to calculate the mean-expression and noise profiles of genes regulated by both sRNAs and microRNAs. We find that despite very different molecular machinery and modes of action (catalytic vs stoichiometric), the mean expression levels and noise profiles of microRNA-regulated genes are almost identical to genes regulated by prokaryotic sRNAs. MicroRNAs suppress noise when proteins are expressed at low levels but substantially increase noise at intermediate and high expression levels. This suggests that microRNAs and sRNAs may represent an example of convergent evolution. We extend our model to study crosstalk between multiple mRNAs that are regulated by a single microRNA and show that noise is a sensitive measure of microRNA-mediated interaction between mRNAs. This suggests a new experimental strategy for uncovering the microRNA-mRNA interactions and testing the competing endogenous RNA (ceRNA) hypothesis.

Author Summary

MicroRNAs and small non-coding RNAs (sRNAs) are small RNA molecules that post-transcriptionally regulate target genes by binding to messenger RNAs (mRNA) and preventing translation. MicroRNAs are found in most eukaryotes and interact with mRNA catalytically (*i.e* a single microRNA molecule regulates multiple mRNA molecules). In contrast, sRNAs are usually found in prokaryotes and regulate target mRNAs in a stoichiometric (one-to-one) fashion. We compare and contrast gene regulation by sRNAs and microRNAs using a mix of computational and analytic techniques. Despite the aforementioned differences, we find that both mean expression levels as well as the intrinsic noise profiles of genes regulated by microRNAs and sRNAs are almost identical. Thus, microRNAs and sRNAs may repre-

sent an example of convergent evolution. We show that gene expression noise is extremely sensitive to microRNA-mediated interactions between mRNAs. This suggests a new experimental strategy for testing the competing endogenous RNA (ceRNA) hypothesis which posits that microRNAs play a key role in mediating interactions between mRNAs in eukaryotes.

Introduction

MicroRNAs are short sequences of RNA (~ 22 base pairs) that post-transcriptionally regulate gene expression in eukaryotes by destabilizing target mRNAs [1, 2]. Since their discovery almost two decades ago [3], there has been a steady increase in the number of discovered microRNAs. MicroRNAs play an important role in many biological processes, including animal development [4, 5], tumor suppression [6, 7], synaptic development [8, 9], programmed cell death [10, 11], and hematopoietic cell fate decisions [12, 13]. In prokaryotes, an analogous role is played by an important class of small non-coding RNAs (antisense sRNAs) that also act post-transcriptionally to negatively regulate proteins. These 100 bp antisense sRNAs prevent translation by binding to the target mRNAs in a process mediated by the RNA chaperone Hfq [14–16].

While both microRNAs and sRNAs play similar functional roles, they act by very different mechanisms [17]. Eukaryotic MicroRNAs undergo extensive post-processing and are eventually incorporated into the RISC assembly [18, 19]. The RISC complex containing the activated microRNA binds mRNAs and targets them for degradation. A single RISC complex molecule can degrade multiple mRNA molecules suggesting that microRNAs act catalytically to repress protein production. In contrast, both the mRNAs and sRNAs are degraded when bound to Hfq [15, 16] suggesting that prokaryotic sRNAs, unlike their eukaryotic counterparts, act stoichiometrically on their targets.

While stoichiometric regulation has been extensively studied theoretically and experimentally [20–26], there exist relatively little work on understanding microRNAs and other catalytic non-coding RNAs [27–29]. Both theoretical calculations (see Supporting Information of [20]) and quantitative experiments [27] indicate that mean protein-expression of microRNA regulated genes follows a linear-threshold behavior similar to that of sRNAs in prokaryotes. In contrast, it was argued in [29] that the intrinsic noise properties of catalytic sRNAs/microRNAs differ significantly from sRNAs. Presently, it is unclear how to reconcile these results and it raises the natural question of how the differing mechanisms employed by sRNAs and microRNAs affect the intrinsic noise profiles of regulated genes.

To answer this question, we used a generalized model of gene regulation by non-coding RNAs to calculate the mean expression and intrinsic noise of regulated proteins. Our model is based on stochastic mass-action kinetics with tunable parameters that allow us to vary the mode of action from stoichiometric interactions such as those in prokaryotic sRNAs to highly catalytic interactions that mimic eukaryotic microRNAs. Contrary to [29], we show that in many parameter regimes the intrinsic noise properties of microRNAs and sRNAs are qualitatively similar. Finally, motivated by the renewed interest in the competing endogenous RNA (ceRNA) hypothesis which suggests that microRNAs induce an extensive mRNA-interaction network, we extended our model to consider the case where a microRNA regulates multiple mRNAs. We calculate the intrinsic noise for these models and show that noise is an extremely sensitive measure of cross talk between mRNAs. This suggests a new experimental strategy for uncovering microRNA-induced mRNA interactions.

Results

Model description

Here we propose a model of gene regulation by non-coding RNAs based on mass-action kinetics [29]. A schematic of the model is shown in Figure 1A. Our model has four species: mRNA molecules denoted

by m , functional non-coding RNAs denoted by s , mRNA- noncoding RNA complexes denoted by c , and the number of expressed proteins denoted by p . We note that s can represent either the number of prokaryotic small RNAs or the number of functional microRNAs found within the RISC complex. mRNA molecules are transcribed at the rates α_m and translated at a rate α_p . In the absence of regulation, mRNAs degrade at a rate τ_m^{-1} . α_s represents the mean rate at which functional non-coding RNAs are formed. For prokaryotic sRNAs, this is simply the transcription rate of sRNAs. For microRNAs, this is an aggregate rate that accounts for the complicated kinetics of transcription and assembly into the RISC complex. mRNAs and noncoding RNAs form complexes c at a rate μ and disassociate at rate γ . Importantly, the complexes can also be degraded at a rate τ_c^{-1} . This degradation can be actively regulated or conversely stem from dilution due to cell division.

Once mRNAs bind noncoding RNA and form a complex c , they can no longer be translated, resulting in decreased protein expression. To account for the differences between microRNAs and sRNAs, we have an additional parameter β which measures the “recycling rate” of noncoding RNA. When $\beta \gg \tau_c^{-1}, \gamma$, the non-coding RNAs function catalytically with a single non-coding RNA molecule able to bind and degrade multiple mRNA molecules. This is a good model for gene regulation by microRNA in eukaryotes. In contrast, when $\beta \ll \tau_c^{-1}$ the noncoding RNAs function stoichiometrically. In particular, for prokaryotic sRNAs it is commonly believed that $\beta = 0$ and no recycling of noncoding RNAs takes place. Finally, we note that the ratio of β to γ is a measure of how much of the gene regulation happens through nucleolytic cleavage as opposed to translational repression.

We use two different approaches to mathematically model gene regulation by non-coding RNAs. We calculate both the mean expression levels as well as “intrinsic noise” due to stochasticity in the underlying biochemical reactions. First, we perform simple Monte-Carlo simulations for the reaction scheme outline above using the Gillespie algorithm [30]. While these simulations are exact, this method is computationally intensive making it difficult to systematically explore how noise properties depend on kinetic parameters. For this reason, we use a second approach based on linear noise approximation (LNA) [22, 31, 32]. The LNA approximates the exact master equations using Langevin equations that correctly reproduce means and variances. This allows us to derive analytical expressions for the noise and systematically explore how gene expression noise depends on parameters.

Within the LNA, we can mathematically represent our model using the equations

$$\begin{aligned}\frac{ds}{dt} &= \alpha_s - \tau_s^{-1}s + \beta c + \gamma c - \mu m s + \eta_s + \eta_\beta + \eta_\gamma - \eta_\mu \\ \frac{dm}{dt} &= \alpha_m - \tau_m^{-1}m + \gamma c - \mu m s + \eta_m + \eta_\gamma - \eta_\mu \\ \frac{dc}{dt} &= \mu m s - \beta c - \gamma c - \tau_c^{-1}c + \eta_c - \eta_\beta - \eta_\gamma + \eta_\mu \\ \frac{dp}{dt} &= \alpha_p m - \tau_p^{-1}p + \eta_p\end{aligned}\tag{1}$$

with η_s, η_m, η_c , and η_p , being the noise in the the birth-death processes of non-coding RNAs, mRNAs, complex, and proteins respectively, η_μ the binding noise, η_γ the unbinding noise, and η_β the RNA recycling noise. The variance of these noise terms is given by

$$\begin{aligned}\langle \eta_i(t) \rangle &= 0, \quad i = s, \beta, \gamma, \mu, m, c, p \\ \langle \eta_i(t) \eta_j(t') \rangle &= \delta_{ij} \sigma_i^2 \delta(t - t')\end{aligned}\tag{2}$$

where δ_{ij} is the Kronecker delta and

$$\begin{aligned}\sigma_s^2 &= \alpha_s + \tau_s^{-1}\bar{s}, \quad \sigma_\beta^2 = \beta\bar{c}, \quad \sigma_\gamma^2 = \gamma\bar{c}, \quad \sigma_\mu^2 = \mu\bar{m}\bar{s}, \\ \sigma_m^2 &= \alpha_m + \tau_m^{-1}\bar{m}, \quad \sigma_c^2 = \tau_c^{-1}\bar{c}, \quad \sigma_p^2 = \alpha_p\bar{m} + \tau_p^{-1}\bar{p},\end{aligned}\tag{3}$$

with \bar{n} the steady-state concentration of species n . We modeled each interaction in Figure 1A as an independent Poisson process with a mean rate given by mass-action kinetics [31]. Each Langevin term is an independent process. Care must be taken to ensure the correct sign in the cross-correlations [32]. In the remainder of the paper, analytical results from the linear noise approximation are shown along with simulations using the Gillespie algorithm [30]. Both methods are in good agreement.

Mean Expression Levels

We begin by deriving the steady-state response of our system by setting the time-derivatives of the left hand side of equation (1) to zero. Denoting the steady-state concentration of a species n by \bar{n} , we find that

$$\begin{aligned}\bar{s} &= \frac{\phi\alpha_s - \alpha_m - \lambda + \sqrt{(\phi\alpha_s - \alpha_m - \lambda)^2 + 4\lambda\phi\alpha_s}}{2\phi\tau_s^{-1}} \\ \bar{m} &= \frac{\alpha_m - \phi\alpha_s - \lambda + \sqrt{(\alpha_m - \phi\alpha_s - \lambda)^2 + 4\lambda\alpha_m}}{2\tau_m^{-1}} \\ \bar{c} &= \mu\bar{m}\bar{s}\tau_{cR} \\ \bar{p} &= \alpha_p\tau_p\bar{m}\end{aligned}\tag{4}$$

where we have defined the quantities

$$\phi = 1 + \beta\tau_c,\tag{5}$$

$$\lambda = \frac{1}{\mu\tau_m\tau_sq},\tag{6}$$

and

$$\tau_{cR}^{-1} = \beta + \gamma + \tau_c^{-1}\tag{7}$$

$$q = \frac{\tau_c^{-1}}{\tau_{cR}^{-1}} = \frac{\tau_c^{-1}}{\beta + \gamma + \tau_c^{-1}}\tag{8}$$

For all values of the parameters, the steady-state protein levels exhibit a linear-threshold behavior in mean protein concentration (Figure 1B) with the threshold set by the condition $\alpha_m \approx \phi\alpha_s$. Such linear-threshold behavior in mean protein production has been extensively studied in the context of non-coding RNAs and it is useful to divide gene expression into three distinct regimes [20, 22, 27]. In the repressed regime ($\alpha_m \ll \phi\alpha_s$) non-coding RNAs almost always bind mRNAs and prevent translation. In contrast, in the expressing regime ($\alpha_m \gg \phi\alpha_s$) there are many more mRNAs than noncoding-RNAs resulting in protein production that varies linearly with α_m . Finally, there is a crossover between these two behaviors when $\alpha_m \approx \phi\alpha_s$.

The factor ϕ multiplying α_s renormalizes the transcription rate of the non-coding RNA to account for the fact that single non-coding RNA can degrade multiple mRNAs. To see this note that $\phi^{-1} = \tau_c^{-1}/(\beta + \tau_c^{-1})$ is just the probability that a non-coding RNA is degraded when it is bound to an mRNA in a complex under the assumption complexes do not disassociate ($\gamma \ll \beta$ or $\gamma \ll \tau_c^{-1}$). Thus, we can think of ϕ as the average number of mRNAs that a non-coding RNA will degrade before it is itself degraded. As expected, when $\beta = 0$, $\phi = 1$ and these results reduce to those derived for prokaryotic sRNAs [20, 22]. Overall, this linear threshold behavior with variable rescaling is consistent with recent experiments on microRNAs [27].

Intrinsic Noise

Gene regulation is inherently stochastic due to the small number of molecules present in the cell. Noise has important phenotypic consequences for a variety of biological phenomena [33] and for this reason it is important to characterize the intrinsic noise properties of non-coding RNAs. As is usual, we define the intrinsic noise as the variance in protein level divided by mean protein level squared, $\frac{\langle (p - \langle p \rangle)^2 \rangle}{\langle p \rangle^2}$, where brackets represent steady state time averaging. This is a measure of relative protein fluctuations compared to their mean. Study of noise in bacterial sRNAs shows a peak in the crossover regime that emerges as a result of competition between mRNA and sRNA [22]. The switch-like behavior of the system, due to its linear-threshold nature, results in an amplification of any fluctuations in the mRNA level that is in excess of the mRNAs bound to noncoding-RNA. As a result, noise is increased in the crossover regime. We observe a similar qualitative behavior for noise in all parameter regimes of our model. Figure 2 is a plot of protein noise as a function of mean protein concentration for catalytic and stoichiometric interactions, showing a peak at protein levels that correspond to the crossover regime in both cases. The height of the peak increases with the binding affinity μ of a non-coding RNA for its target mRNA. On a whole, the noise profile of catalytic and non-catalytic genes is remarkably similar. There are however slight differences. The peak is slightly higher and occurs at a slightly larger mean protein level for catalytic non-coding RNAs. This suggests that the underlying reason for the noise peak is not the enzymatic mode of action of non-coding RNAs, but the fact that the number of mRNAs and the number of non-coding RNAs (appropriately normalized by ϕ) are almost equal.

To better understand the effect of catalytic interaction on noise, we calculated protein noise versus α_m for different values of β . Figure 3A shows the results for Gillespie simulation and linear noise approximation. To meaningfully compare noise between stoichiometric (i.e. prokaryotic sRNA, $\beta = 0$) and catalytic (i.e. eukaryotic microRNA, $\beta \gg \tau_c^{-1}, \gamma$) we need to compare noise at the same mean protein level. Figure 3B shows this comparison for noise in the repressed regime as a function of β with α_m chosen such as to keep mean protein concentration constant. As can be seen, the noise decreases from its original value in the stoichiometric regime ($\beta = 0$) to a slightly lower value as β is slightly increased ($\beta \approx 2$) and any further increase in β does not affect noise dramatically. We also plotted the maximum of the noise peak in the crossover regime using both linear noise approximation and Gillespie algorithm (Figure 3C). Notice that the peak height initially increases as a function of β and then reaches a plateau for large β upon entering the catalytic regime.

Including transcriptional bursting

Experimental evidence suggests that mRNAs are often produced in bursts [34]. Transcriptional bursting represents another important source of stochasticity that was ignored in the analysis presented above. We can extend the model presented above to incorporate transcriptional bursting by considering the case where the genes encoding for mRNAs can be in two distinct states: an “on” state where mRNAs can be transcribed and an “off” state where transcription is not possible. For example, in eukaryotes the two states may correspond to whether the chromatin is condensed or not [35]. To model transcriptional bursting we replace the equation for mRNA production in Eq. (1) by the pair of equations

$$\begin{aligned} \frac{dg}{dt} &= k_-(1 - g) - k_+g + \eta_g \\ \frac{dm}{dt} &= \alpha_m^{on} g - \tau_m^{-1}m + \gamma c - \mu ms + \eta_m + \eta_\gamma - \eta_\mu \end{aligned} \quad (9)$$

where the probability that a gene is in the on state is denoted by g , $k_-(k_+)$ are the on(off) rate of the gene, and α_m^{on} is the maximum mRNA transcription rate. The mean probability of the gene being on is given by $\bar{g} = k_-/(k_- + k_+)$ [22]. The gene noise η_g is Gaussian white noise with mean zero and variance

given by $\langle \eta_g(t) \eta_g(t') \rangle = 2k_+ \bar{g} \delta(t-t')$. The mRNA noise η_m is now $\langle \eta_m(t) \eta_m(t') \rangle = (\alpha_m^{on} \bar{g} + \tau_m^{-1} \bar{m}) \delta(t-t')$. All other equations remain the same as before.

In the analysis that follows, we assume that k_- is fixed but k_+ can vary. This corresponds to the biological situation where a gene is regulated by a repressor that can turn the gene off. To compare noise for different values of β , we choose k_+ so that the mean protein levels remain constant. Furthermore, we concentrate on the repressed regime (see Figure 4). Here noise decreases slightly as β is increased which is very similar to the case without bursting as was shown in Figure 3B. This means that noise in the repressed regime is insensitive to bursting regardless of how catalytic the interaction is.

Asymptotic Formulas for Noise

Expressed

To gain further insight, we have derived asymptotic formulas for the noise in the repressed and expressing regime. In the expressing regime with large mRNA transcription, we define the small parameter $\epsilon \equiv \frac{\lambda}{\alpha_m - \phi \alpha_s}$ with λ given by Eq. 6. In the expressing regime, $\epsilon \ll 1$, and the steady-state expression levels of mRNA and non-coding RNAs take the form

$$\bar{m} \sim \frac{1}{q\mu\tau_s} \frac{1}{\epsilon}, \quad \bar{s} \sim \alpha_s \tau_s \epsilon \quad (10)$$

with q given by Eq. 8. Furthermore, the protein noise in this regime is identical to that of a transcriptionally regulated gene and is given by

$$\frac{\langle (\delta p)^2 \rangle}{\bar{p}^2} \sim \frac{1 + b \frac{\tau_p}{\tau_p + \tau_m}}{\bar{p}} \sim \frac{1 + b}{\bar{p}}$$

where $b \equiv \alpha_p \tau_m$ (see Materials and Methods) and it is assumed that mRNAs degrade much faster than proteins ($\tau_p \gg \tau_m$). The quantity b is often called the ‘burst size’ and measures the average number of proteins made from a single mRNA molecule [36]. This result shows that in the expressing regime the protein noise has no β dependence, and thus the protein noise of post-transcriptional and transcriptional regulation are identical in the asymptotic limit.

Repressed

In the repressed regime, we now have $\epsilon \equiv \frac{\lambda}{\phi \alpha_s - \alpha_m} \ll 1$ and the average number of mRNA and non-coding RNA molecules is given by

$$\bar{m} \sim \alpha_m \tau_m \epsilon \quad (11)$$

$$\bar{s} \sim \frac{1}{q\mu\tau_m\phi\epsilon}$$

where q

$$q \equiv \frac{\tau_c^{-1}}{\beta + \gamma + \tau_c^{-1}},$$

is the probability that a complex is degraded. The noise in this regime is given by

$$\frac{\langle (\delta p)^2 \rangle}{\bar{p}^2} \sim \frac{1 + \zeta b \epsilon}{\bar{p}} \quad (12)$$

where ζ is a constant that is dependent on both β and γ (see Materials and Methods). When β or $\tau_c^{-1} \gg \gamma$ it can be shown that $\zeta \approx 1$. Note that this condition includes the catalytic regime with

$\beta \gg \gamma, \tau_c^{-1}$ as well as the stoichiometric regime $\beta = 0, \tau_c^{-1} \gg \gamma$ (see Methods). Thus, we conclude that in the repressed regime, non-coding RNAs reduce noise. In particular, proteins are now produced from mRNAs in a smaller burst with typical size given by $\epsilon \zeta b \ll b$. Since the burst size is just the average number of proteins made from an mRNA, $b = \alpha_b \tau_m$, we can equivalently interpret this results as changing the effective lifetime of the mRNA molecules from τ_m to $\epsilon \zeta \tau_m$ [20, 22].

Scaling Behavior Near Crossover Regime

Our analysis show that the width of noise peak at the crossover regime is independent of recycling ratio β . To understand this behavior better, we studied the crossover regime for different parameter values using linear noise approximation. Figure 5 shows the protein noise and protein mean as a function of $\frac{\alpha_m}{\phi \alpha_s}$ after rescaling both by their value at the the crossover ($\alpha_m = \phi \alpha_s$) for various recycling ratios. Notice that for $\gamma \ll \tau_c^{-1}$ these normalized plots of protein mean and noise show an approximate data collapse. As γ is increased, this scaling behavior breaks down (Figure 5). The collapse of data for mean protein can be analytically derived given the fact that mean protein is only dependent on β, γ, τ_c through the combination $q\phi \equiv \frac{\beta + \tau_c^{-1}}{\tau_c^{-1} + \beta + \gamma}$.

This scaling of the mean protein number can be better understood if we define $x \equiv \frac{\alpha_m}{\phi \alpha_s}$ and define $p(x)$ as the mean number of proteins corresponding to this value. Dividing this quantity by its value at $x = 1$ and using equation (4) results in the following equation for the normalized mean protein:

$$\frac{p(x)}{p(1)} = \frac{x - 1 - \frac{\nu}{q\phi} + \sqrt{\left(x - 1 - \frac{\nu}{q\phi}\right)^2 + 4\frac{\nu}{q\phi}x}}{-\frac{\nu}{q\phi} + \sqrt{\left(\frac{\nu}{q\phi}\right)^2 + 4\frac{\nu}{q\phi}}}. \quad (13)$$

where $\nu \equiv \frac{\tau_m^{-1} \tau_s^{-1}}{\mu \alpha_s}$ is a constant with no dependence on β, γ, τ_c . Thus, the normalized mean protein number depends on β, γ, τ_c only through the combination $q\phi \equiv \frac{\beta + \tau_c^{-1}}{\tau_c^{-1} + \beta + \gamma}$. The parameter $q\phi$ is the probability a complex will disassociate. In the limit $\gamma \ll \beta, \tau_c^{-1}$ this parameter will be equal to 1 and the scaled mean protein level becomes independent of β causing the curves for different β to collapse on top of each other (figure 5A). It is also interesting to note that any other condition on the parameters that removes the dependence of $q\phi$ on β will also have the same effect (e.g. $\gamma, \beta \ll \tau_c^{-1}$). Somewhat more surprisingly, the noise also shows an approximate data collapse. We suspect that this collapse is likely indicative of universality near the crossover between the repressed and expressing regimes.

mRNA Crosstalk and the ceRNA Hypothesis

Recently, the competing endogenous RNA (ceRNA) hypothesis proposed that microRNAs may play a crucial role in the cell in global gene regulation by inducing interactions between mRNA species [37]. The central mechanism underlying the ceRNA hypothesis is the idea that mRNA species may have interactions amongst themselves that are not direct but are instead indirect and mediated by competition and depletion of shared microRNA pools. The hypothesis is that these indirect mRNA interactions results in a biologically important mRNA network. However, the breadth and strength of microRNA induced interactions in eukaryotic genomes is still not well understood. For this reason, it is crucial to develop new strategies for measuring microRNA induced interactions between commonly regulated mRNAs. To this end, we asked whether the noise profile of regulated mRNAs could be used to uncover microRNA induced interactions. As a first step, we studied the simplified case where two different mRNA species are regulated by a single microRNA and compete over a common pool of microRNAs and focused on the

effect of microRNA-induced crosstalk between mRNA species on the noise properties of regulated genes. The results presented here can be easily generalized to the case of many mRNAs interacting with many microRNAs.

A schematic of our simplified model is shown in Figure 6A. Two species of mRNAs are regulated by a common microRNA. Notice that the mRNAs do not directly interact in the model and all interactions are indirectly induced by the shared microRNA pool. We can represent this using a straight-forward generalization of the model considered earlier

$$\begin{aligned}
\frac{ds}{dt} &= \alpha_s - \tau_s^{-1}s + \beta_1c_1 + \beta_2c_2 + \gamma_1c_1 + \gamma_2c_2 - (\mu_1m_1 + \mu_2m_2)s + \eta_s + \eta_{\beta_1} + \eta_{\beta_2} + \eta_{\gamma_1} + \eta_{\gamma_2} - \eta_{\mu_1} - \eta_{\mu_2} \\
\frac{dm_i}{dt} &= \alpha_{m_i} - \tau_{m_i}^{-1}m_i + \gamma_i c_i - \mu_i m_i s + \eta_{m_i} + \eta_{\gamma_i} - \eta_{\mu_i} \\
\frac{dc_i}{dt} &= \mu_i m_i s - \beta_i c_i - \gamma_i c_i - \tau_{c_i}^{-1}c_i + \eta_{c_i} - \eta_{\beta_i} - \eta_{\gamma_i} + \eta_{\mu_i} \\
\frac{dp}{dt} &= \alpha_p m_1 - \tau_p^{-1}p + \eta_p
\end{aligned} \tag{14}$$

with

$$\begin{aligned}
\langle \eta_j(t) \rangle &= 0, \quad j = s, \beta_i, \gamma_i, \mu_i, m_i, c_i, p, \quad i = 1, 2 \\
\langle \eta_i(t) \eta_j(t') \rangle &= \delta_{ij} \sigma_i^2 \delta(t - t')
\end{aligned} \tag{15}$$

and variances reflecting the Poisson nature of interactions:

$$\begin{aligned}
\sigma_s^2 &= \alpha_s + \tau_s^{-1}\bar{s}, \quad \sigma_{\beta_i}^2 = \beta_i \bar{c}_i, \quad \sigma_{\gamma_i}^2 = \gamma_i \bar{c}_i, \quad \sigma_{\mu_i}^2 = \mu_i \bar{m}_i \bar{s} \\
\sigma_{m_i}^2 &= \alpha_{m_i} + \tau_{m_i}^{-1}\bar{m}_i, \quad \sigma_{c_i}^2 = \tau_{c_i}^{-1}\bar{c}_i, \quad \sigma_p^2 = \alpha_p \bar{m}_1 + \tau_p^{-1}\bar{p}, \quad i = 1, 2
\end{aligned} \tag{16}$$

The binding of microRNAs to mRNAs is controlled by the interaction rates $\mu_{1,2}$. These rates reflect the binding affinity of microRNAs for the two mRNA species. We asked how protein noise and means change as we vary the transcription rates, $\alpha_{m_{1,2}}$, of the mRNAs. The remaining parameters are assumed to be the same for both mRNAs and we have suppressed the indices on these parameters. MicroRNAs function catalytically and we focus on the parameter regime $\beta \gg \tau_c^{-1}, \gamma$. Figure 6B and C shows the mean protein levels, \bar{p}_1 , and intrinsic noise of protein species 1 as a function of the transcription rates of the two mRNA genes, $\alpha_{m_{1,2}}$, for the case of equal binding affinities ($\mu_1 = \mu_2$). Notice there is a peak in the noise similar to the single-species case. Once again there is good agreement between simulation and analytic calculations based on Langevin noise. We also examined the case where the mRNAs have different binding affinities for the microRNA, $\mu_1 = 0.2, \mu_2 = 2$. This results in an asymmetry in the noise peak but does not change the major qualitative features of our results (see Figure 6D and E)

As in the single-mRNA species case, the behavior of the system can be divided into regimes depending on whether the combined transcription rate of both mRNA species is bigger or smaller than the normalized sRNA transcription rates. The crossover regime, $\alpha_{m_1} + \alpha_{m_2} \approx \phi \alpha_s$ splits the transcription rate space $(\alpha_{m_1}, \alpha_{m_2})$ into an expressing regime, $\alpha_{m_1} + \alpha_{m_2} > \phi \alpha_s$, and a repressing regime, $\alpha_{m_1} + \alpha_{m_2} < \phi \alpha_s$. Here, similar to the single species case we see a sharp peak in the noise at the crossover regime. Since the cross-over regime depends on the total transcription rate, $\alpha_{m_1} + \alpha_{m_2}$, this suggests a general strategy for uncovering microRNA interactions based on ‘‘intrinsic noise spectroscopy’’. To see if an mRNA, say species 2, interacts with an mRNA species 1 that is repressed by a microRNA ($\alpha_{m_1} \ll \phi \alpha_s$), we can monitor the intrinsic noise of protein 1 (i.e. a peak) as we vary α_{m_2} while keeping α_{m_1} fixed. If mRNA species 2 interacts with the microRNA, there will be a dramatic signature of the interaction in the intrinsic noise of protein 1. In contrast, there are no dramatic features in the mean levels of proteins.

Discussion

In this work, we studied gene regulation by non-coding RNAs. Whereas gene regulation by prokaryotic sRNAs has been extensively studied [20–26], there exist relatively few models of gene regulation by catalytic microRNAs [27–29]. Here, we used a simple kinetic model to study both the mean expression levels and intrinsic noise properties of post-transcriptional regulation by non-coding RNAs. Using a single parameter, our model interpolates between the stoichiometric behavior of prokaryotic sRNAs and the catalytic behavior characteristic of eukaryotic microRNAs. We found that both sRNAs and microRNAs exhibit a linear threshold behavior, with expressing and repressed regimes separated by a crossover regime. At the crossover, the mRNA transcription rate roughly equals the product of the non-coding-RNA transcription rate and the average number of mRNA molecules degraded by a single non-coding RNA molecule. In all cases, the intrinsic noise was smaller in the repressed regime and showed a sharp peak in the crossover regime. We found that for most parameter regimes, the intrinsic noise in the crossover regime shows an approximate data collapse, giving hints that there may be universal behavior near the transition from the repressed to expressing regime. We then generalized our calculations to study between two mRNAs regulated by a single microRNA. We found that the intrinsic noise is an extremely sensitive measure of microRNA induced interactions between mRNAs.

Our results for the mean expression profile is consistent with recent experimental studies [27]. However, our conclusions about the intrinsic noise profiles of catalytic non-coding RNAs are different from Hao *et al.* [29]. They concluded that the intrinsic noise profiles of catalytic and stoichiometric differed significantly. The reason for this discrepancy is that Hao *et al.* did not normalize the sRNA transcription rate α_s by the recycling rate ϕ . Consequently, they compared the cross-over regions of sRNA regulated genes to repressed regions of microRNAs. As shown above, after making this normalization there is extensive data collapse of intrinsic noise profiles for both stoichiometric and catalytic genes.

One of the striking results of our calculation is the similarity between sRNA-regulated and microRNA-regulated genes. This similarity is somewhat surprising given that microRNAs and sRNAs are found in different kingdoms (prokaryotes versus eukaryotes) and utilize distinct biochemistry and molecular machinery. Eukaryotic microRNAs use complicated nuclear machinery to export microRNA into cytoplasm and bring it to mature state by incorporating the RNA strand into the RISC complex. In contrast, prokaryotic sRNAs undergo relatively little post-processing and bind mRNAs via the chaperone protein Hfq. Given these extensive differences, the similarity between the expression characteristics of microRNA-regulated and sRNA regulated genes are suggestive of convergent evolution.

Our calculations show that mRNAs regulated by catalytic non-coding RNAs have large peaks in the intrinsic noise. This differs significantly from results that would be derived from more traditional treatments of catalytic interactions based on the Michaelis-Menten or Hill equations. The underlying reason for this difference is that traditionally, the Michaelis-Menten equations are derived under the assumption that substrates of enzymes are in excess compared to the enzymes themselves. In contrast, here we are interested in the case where the number of enzymes (microRNAs) and number of substrates (mRNAs) are comparable. This accounts for the sharp peak in noise observed in the crossover regime in our models that is absent in traditional treatments of enzyme kinetics.

Our calculations also suggest a strategy for testing ceRNA hypothesis [37], which posits that microRNAs induce extensive interactions between mRNA molecules. Our calculations suggest that protein noise is a more sensitive measure of the competition between the two species than mean levels. Thus, it may be easier to uncover interactions between mRNA by measuring changes in the noise of regulated genes. We suspect that this will be true even when we generalize our calculations to consider the case where n mRNA species compete over the same pool of microRNAs. In this case, we hypothesize that there would still be a sharp peak in the intrinsic noise when the total transcription rate of all mRNAs equals the appropriately normalized transcription rate of microRNAs. In the future, it will be interesting to analyze these more complicated models in greater detail.

Materials and Methods

Single Species Linear noise approximation

Linearization of equation 1 results in:

$$\frac{d}{dt} \begin{bmatrix} \delta s \\ \delta m \\ \delta c \end{bmatrix} = \underbrace{\begin{bmatrix} -\tau_{sR}^{-1} & -\mu\bar{s} & \beta + \gamma \\ -\mu\bar{m} & -\tau_{mR}^{-1} & \gamma \\ \mu\bar{n} & \mu\bar{s} & -\tau_{cR}^{-1} \end{bmatrix}}_A \begin{bmatrix} \delta s \\ \delta m \\ \delta c \end{bmatrix} + \begin{bmatrix} \eta_s + \eta_\beta + \eta_\gamma - \eta_\mu \\ \eta_m + \eta_\gamma - \eta_\mu \\ \eta_c - \eta_\beta - \eta_\gamma + \eta_\mu \end{bmatrix}$$

where we have named the transfer matrix A and we define the renormalized lifetimes:

$$\tau_{sR}^{-1} = \tau_s^{-1} + \mu\bar{m}, \quad \tau_{mR}^{-1} = \tau_m^{-1} + \mu\bar{s}, \quad \tau_{cR}^{-1} = \tau_c^{-1} + \beta + \gamma \quad (17)$$

To find the solution of this equation we apply a Fourier transform and solve the resulting equation:

$$\begin{bmatrix} \widetilde{\delta s} \\ \widetilde{\delta m} \\ \widetilde{\delta c} \end{bmatrix} = \begin{bmatrix} i\omega + \tau_{sR}^{-1} & \mu\bar{s} & -\beta - \gamma \\ \mu\bar{m} & i\omega + \tau_{mR}^{-1} & -\gamma \\ -\mu\bar{n} & -\mu\bar{s} & i\omega + \tau_{cR}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\eta}_s + \tilde{\eta}_\beta + \tilde{\eta}_\gamma - \tilde{\eta}_\mu \\ \tilde{\eta}_m + \tilde{\eta}_\gamma - \tilde{\eta}_\mu \\ \tilde{\eta}_c - \tilde{\eta}_\beta - \tilde{\eta}_\gamma + \tilde{\eta}_\mu \end{bmatrix}$$

where tilde denotes Fourier transform. So we have:

$$\widetilde{\delta m} = \frac{1}{(i\omega - \lambda_1)(i\omega - \lambda_2)(i\omega - \lambda_3)} \begin{bmatrix} F(\omega) & G(\omega) & H(\omega) \end{bmatrix} \begin{bmatrix} \tilde{\eta}_s + \tilde{\eta}_\beta + \tilde{\eta}_\gamma - \tilde{\eta}_\mu \\ \tilde{\eta}_m + \tilde{\eta}_\gamma - \tilde{\eta}_\mu \\ \tilde{\eta}_c - \tilde{\eta}_\beta - \tilde{\eta}_\gamma + \tilde{\eta}_\mu \end{bmatrix}$$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of A and:

$$F(\omega) = \gamma\mu\bar{n} - \mu\bar{m}(i\omega + \tau_{cR}^{-1}) \quad (18)$$

$$G(\omega) = (i\omega + \tau_{sR}^{-1})(i\omega + \tau_{cR}^{-1}) - (\beta + \gamma)\mu\bar{m} \quad (19)$$

$$H(\omega) = -(\beta + \gamma)\mu\bar{m} + \gamma(i\omega + \tau_{sR}^{-1}) \quad (20)$$

or

$$\widetilde{\delta m} = \frac{F(\omega)\tilde{\eta}_s + [F(\omega) - H(\omega)]\tilde{\eta}_\beta + [F(\omega) + G(\omega) - H(\omega)](\tilde{\eta}_\gamma - \tilde{\eta}_\mu) + G(\omega)\tilde{\eta}_m + H(\omega)\tilde{\eta}_c}{(i\omega - \lambda_1)(i\omega - \lambda_2)(i\omega - \lambda_3)}$$

Since $\widetilde{\delta p} = \frac{\tilde{\eta}_p + \alpha_p \widetilde{\delta m}}{i\omega + \tau_p^{-1}}$, we have

$$\langle (\delta p)^2 \rangle = \frac{\tau_p}{2} \sigma_p^2 + \int \frac{d\omega}{2\pi} \frac{\alpha_p^2}{\omega^2 + \tau_p^{-2}} \times \quad (21)$$

$$\overbrace{\frac{|F(\omega)|^2 \sigma_s^2 + |F(\omega) - H(\omega)|^2 \sigma_\beta^2 + |F(\omega) + G(\omega) - H(\omega)|^2 (\sigma_\gamma^2 + \sigma_\mu^2) + |G(\omega)|^2 \sigma_m^2 + |H(\omega)|^2 \sigma_c^2}{(\omega^2 + \lambda_1^2)(\omega^2 + \lambda_2^2)(\omega^2 + \lambda_3^2)}}^Q$$

We can rewrite Q as:

$$Q \equiv X\omega^4 + Y\omega^2 + Z \quad (22)$$

with

$$\begin{aligned}
X &= (\gamma\tau_{cR} + 1)\mu\bar{m}\bar{s} + \alpha_m + \tau_m^{-1}\bar{m} \\
Y &= (\mu\bar{m})^2(\alpha_s + \tau_s^{-1}\bar{s}) + (\mu\bar{m} + \gamma)^2\beta\tau_{cR}\mu\bar{m}\bar{s} \\
&\quad + (\tau_s^{-1} + \tau_c^{-1} + \beta)^2(\gamma\tau_{cR} + 1)\mu\bar{m}\bar{s} \\
&\quad + [(\tau_{sR}^{-1} - \tau_{cR}^{-1})^2 + 2(\beta + \gamma)\mu\bar{m}]^2(\alpha_m + \tau_m^{-1}\bar{m}) + \gamma^2q\mu\bar{m}\bar{s} \\
Z &= (\gamma - \tau_{cR}^{-1})^2(\mu\bar{m})^2(\alpha_s + \tau_s^{-1}\bar{s}) + (\tau_c^{-1}\mu\bar{m} + \gamma\tau_s^{-1})^2\beta\tau_{cR}\mu\bar{m}\bar{s} \\
&\quad + (\tau_c^{-1} + \beta)^2(\gamma\tau_{cR} + 1)\tau_s^{-2}\mu\bar{m}\bar{s} \\
&\quad + (\tau_s^{-1}\tau_{cR}^{-1} + \tau_c^{-1}\mu\bar{m})^2(\alpha_m + \tau_m^{-1}\bar{m}) + (\beta\mu\bar{m} - \gamma\tau_s^{-1})^2q\mu\bar{m}\bar{s}
\end{aligned} \tag{23}$$

now if we write

$$\langle (\delta p)^2 \rangle = \frac{\tau_p}{2}\sigma_p^2 + \alpha_p^2 \int \frac{d\omega}{2\pi} \frac{X\omega^4 + Y\omega^2 + Z}{(\omega^2 + \lambda_1^2)(\omega^2 + \lambda_2^2)(\omega^2 + \lambda_3^2)(\omega^2 + \tau_p^{-2})} \tag{24}$$

and let $\lambda_4 = -\tau_p^{-1}$ by use of partial fractions and integration we get

$$\langle (\delta p)^2 \rangle = \frac{\tau_p}{2}(\alpha_p\bar{m} + \tau_p^{-1}\bar{p}) + \sum_{i=1}^4 \frac{\alpha_p^2}{2|\lambda_i|} (X\lambda_i^4 - Y\lambda_i^2 + Z) \prod_{j \neq i} \frac{1}{\lambda_j^2 - \lambda_i^2} \tag{25}$$

Single Species Asymptotic calculations

Expressing Regime

In the expressing regime with large mRNA transcription rate we demand $\epsilon \equiv \frac{\lambda}{\alpha_m - \phi\alpha_s} \ll 1$. Expanding \bar{m} and \bar{s} in terms of ϵ we will get:

$$\bar{m} = \frac{1}{q\mu\tau_s} \frac{1}{\epsilon} + O(\epsilon) \tag{26}$$

$$\bar{s} = \alpha_s\tau_s\epsilon + O(\epsilon^2) \tag{27}$$

To find the eigenvalues of A we will expand it in terms of ϵ :

$$A = \begin{bmatrix} -\tau_s^{-1} - \tau_s^{-1}q^{-1}\epsilon^{-1} + O(\epsilon) & -\mu\alpha_s\tau_s\epsilon + O(\epsilon^2) & \beta + \gamma \\ -\tau_s^{-1}q^{-1}\epsilon^{-1} + O(\epsilon) & -\tau_m^{-1} - \mu\alpha_s\tau_s\epsilon + O(\epsilon^2) & \gamma \\ \tau_s^{-1}q^{-1}\epsilon^{-1} + O(\epsilon) & \mu\alpha_s\tau_s\epsilon + O(\epsilon^2) & -\tau_{cR}^{-1} \end{bmatrix}$$

the eigenvalues of this matrix satisfy the following equation:

$$\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0 \tag{28}$$

where

$$a_2 = \tau_s^{-1}q^{-1}\epsilon^{-1} + O(1) \tag{29}$$

$$a_1 = \tau_s^{-1}q^{-1}(\tau_m^{-1} + \tau_c^{-1})\epsilon^{-1} + O(1)$$

$$a_0 = q^{-1}\tau_s^{-1}\tau_m^{-1}\tau_c^{-1}\epsilon^{-1} + O(1)$$

this equation can be analytically solved by expanding λ 's in terms of ϵ , and results in the following:

$$\begin{aligned}\lambda_1 &= -\tau_m^{-1} + O(\epsilon) \\ \lambda_2 &= -\tau_c^{-1} + O(\epsilon) \\ \lambda_3 &= -\frac{\tau_s^{-1}q^{-1}}{\epsilon} + O(1)\end{aligned}\tag{30}$$

which shows one fast mode and two slow modes. Next we will calculate the noise by expanding equation 24 in terms of ϵ , which results in:

$$\begin{aligned}X &\sim \frac{2\lambda}{\epsilon} \\ Y &\sim \frac{2\lambda}{q^2\tau_s^2\epsilon^3} \\ Z &\sim \frac{2\lambda}{q^2\tau_s^2\tau_c^2\epsilon^3}\end{aligned}\tag{31}$$

plugging this result into equation 25 and using the approximation $\tau_p \gg \tau_m$ we will have:

$$\frac{\langle(\delta p)^2\rangle}{\bar{p}^2} \sim \frac{1+b}{\bar{p}}$$

Repressing Regime

In this regime with mRNA transcription very small, we demand $\epsilon \equiv \frac{\lambda}{\phi\alpha_s - \alpha_m} \ll 1$. Expansion of \bar{m} and \bar{s} in terms of ϵ gives:

$$\bar{m} = \alpha_m\tau_m\epsilon + O(\epsilon^2) \quad \bar{s} = \frac{1}{q\mu\tau_m\phi\epsilon} + O(\epsilon)\tag{32}$$

Next we expand A in terms of ϵ :

$$A = \begin{bmatrix} -\tau_s^{-1} - \mu\alpha_m\tau_m\epsilon + O(\epsilon^2) & -\frac{1}{q\tau_m\phi\epsilon} + O(\epsilon) & \beta + \gamma \\ -\mu\alpha_m\tau_m\epsilon + O(\epsilon^2) & -\tau_m^{-1} - \frac{1}{q\tau_m\phi\epsilon} + O(\epsilon) & \gamma \\ \mu\alpha_m\tau_m\epsilon + O(\epsilon^2) & \frac{1}{q\tau_m\phi\epsilon} + O(\epsilon) & -\tau_c^{-1} \end{bmatrix}$$

after some calculation we find the following closed form for eigenvalues of A :

$$\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0\tag{33}$$

where

$$\begin{aligned}a_2 &= \tau_m^{-1}q^{-1}\phi^{-1}\epsilon^{-1} + O(1) \\ a_1 &= \tau_m^{-1}q^{-1}\phi^{-1}(\tau_s^{-1} + \tau_c^{-1} + \beta)\epsilon^{-1} + O(1) \\ a_0 &= \tau_m^{-1}q^{-1}\phi^{-1}\tau_s^{-1}(\tau_c^{-1} + \beta)\epsilon^{-1} + O(1)\end{aligned}\tag{34}$$

this equation can be analytically solved by expanding λ 's in terms of ϵ , and results in the following:

$$\begin{aligned}\lambda_1 &= -\tau_s^{-1} + O(\epsilon) \\ \lambda_2 &= -(\tau_c^{-1} + \beta) + O(\epsilon) \\ \lambda_3 &= -\frac{\tau_m^{-1}q^{-1}\phi^{-1}}{\epsilon} + O(1)\end{aligned}\tag{35}$$

which has one fast mode and two slow modes. Next we calculate the noise by expanding equation 24 in terms of ϵ :

$$\begin{aligned} X &\sim 2\alpha_m q^{-1} \phi^{-1} \\ Y &\sim 2\alpha_m (\tau_s^{-2} \tau_{cR}^{-1} q^{-1} \phi^{-1} + \gamma \tau_s^{-1} + \tau_{cR}^{-2}) \\ Z &\sim 2\alpha_m \tau_s^{-2} \tau_{cR}^{-2} \end{aligned} \quad (36)$$

plugging this result into equation 25 gives:

$$\begin{aligned} \frac{\langle (\delta p)^2 \rangle}{\bar{p}^2} &\sim \frac{1}{\bar{p}} + \frac{b\epsilon}{\bar{p}} \frac{q^2 \phi^2}{2\alpha_m \tau_p} \left(\frac{X\tau_p^{-4} - Y\tau_p^{-2} + Z}{(\tau_s^{-2} - \tau_p^{-2}) ((\tau_c^{-1} + \beta)^2 - \tau_p^{-2}) \tau_p^{-1}} + \frac{X\tau_s^{-4} - Y\tau_s^{-2} + Z}{(\tau_p^{-2} - \tau_s^{-2}) ((\tau_c^{-1} + \beta)^2 - \tau_s^{-2}) \tau_s^{-1}} \right. \\ &\quad \left. + \frac{X(\tau_c^{-1} + \beta)^4 - Y(\tau_c^{-1} + \beta)^2 + Z}{(\tau_p^{-2} - (\tau_c^{-1} + \beta)^2) (\tau_s^{-2} - (\tau_c^{-1} + \beta)^2) (\tau_c^{-1} + \beta)} \right) + O(\epsilon^3) \end{aligned} \quad (37)$$

for $\beta \gg \gamma, \tau_c^{-1}, \tau_s^{-1}, \tau_p^{-1}$, after some calculation we get $\frac{\langle (\delta p)^2 \rangle}{\bar{p}^2} \sim \frac{1}{\bar{p}} (1 + b\epsilon)$.

ceRNA Linear Noise Approximation

For two mRNAs, we linearized equation 15 as:

$$\frac{d\delta\chi}{dt} = A\delta\chi + \xi$$

$$\delta\chi = \begin{bmatrix} \delta s \\ \delta m_1 \\ \delta m_2 \\ \delta c_1 \\ \delta c_2 \\ \delta p_1 \\ \delta p_2 \end{bmatrix}, \quad A = \begin{bmatrix} -\tau_{sR}^{-1} & -\mu_1 \bar{s} & -\mu_2 \bar{s} & \beta_1 + \gamma_1 & \beta_2 + \gamma_2 & 0 & 0 \\ -\mu_1 \bar{m}_1 & -\tau_{mR_1}^{-1} & 0 & \gamma_1 & 0 & 0 & 0 \\ -\mu_2 \bar{m}_2 & 0 & -\tau_{mR_2}^{-1} & 0 & \gamma_2 & 0 & 0 \\ \mu_1 \bar{m}_1 & \mu_1 \bar{s} & 0 & -\tau_{cR_1}^{-1} & 0 & 0 & 0 \\ \mu_2 \bar{m}_2 & 0 & \mu_2 \bar{s} & 0 & -\tau_{cR_2}^{-1} & 0 & 0 \\ 0 & \alpha_{p_1} & 0 & 0 & 0 & -\tau_{p_1}^{-1} & 0 \\ 0 & 0 & \alpha_{p_2} & 0 & 0 & 0 & -\tau_{p_2}^{-1} \end{bmatrix}, \quad \xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \end{bmatrix}$$

with

$$\tau_{sR}^{-1} = \tau_s^{-1} + \mu_1 \bar{m}_1 + \mu_2 \bar{m}_2, \quad \tau_{mR_i}^{-1} = \tau_{m_i}^{-1} + \mu_i \bar{s}, \quad \tau_{cR_i}^{-1} = \tau_{c_i}^{-1} + \gamma_i + \beta_i$$

and:

$$< \xi_i(t) \xi_j(t') > = \Gamma_{ij} \delta(t - t')$$

where

$$\Gamma = \begin{bmatrix} g_1 & g_2 & g_3 & g_4 & g_5 & 0 & 0 \\ g_2 & g_6 & 0 & g_7 & 0 & 0 & 0 \\ g_3 & 0 & g_8 & 0 & g_9 & 0 & 0 \\ g_4 & g_7 & 0 & g_{10} & 0 & 0 & 0 \\ g_5 & 0 & g_9 & 0 & g_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & g_{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g_{13} \end{bmatrix} \quad (38)$$

and

$$\begin{aligned}
g_1 &= \sigma_s^2 + \sigma_{\beta_1}^2 + \sigma_{\gamma_1}^2 - \sigma_{\mu_1}^2 + \sigma_{\beta_2}^2 + \sigma_{\gamma_2}^2 - \sigma_{\mu_2}^2 \\
g_2 &= \sigma_{\gamma_1}^2 + \sigma_{\mu_1}^2 \\
g_3 &= \sigma_{\gamma_2}^2 + \sigma_{\mu_2}^2 \\
g_4 &= -\sigma_{\beta_1}^2 - \sigma_{\gamma_1}^2 - \sigma_{\mu_1}^2 \\
g_5 &= -\sigma_{\beta_2}^2 - \sigma_{\gamma_2}^2 - \sigma_{\mu_2}^2 \\
g_6 &= \sigma_{m_1}^2 + g_2 \\
g_7 &= -g_2 \\
g_8 &= \sigma_{m_2}^2 + g_3 \\
g_9 &= -g_3 \\
g_{10} &= \sigma_{c_1}^2 - g_4 \\
g_{11} &= \sigma_{c_2}^2 - g_5 \\
g_{12} &= \sigma_{p_1}^2 \\
g_{13} &= \sigma_{p_2}^2
\end{aligned} \tag{39}$$

We find corresponding two point correlation functions by use of the following equation [32]:

$$\langle \delta\chi_i \delta\chi_j \rangle = - \sum_{p,q,r,s} B_{ip} B_{jr} \frac{\Gamma_{qs}}{\lambda_p + \lambda_r} B_{pq}^{-1} B_{rs}^{-1}$$

where λ 's are the eigenvalues of A , and B is the matrix of eigenvectors, according to:

$$\sum_j A_{ij} B_{ij} = \lambda_k B_{ik}$$

we saw a good agreement between these results and Gillespie simulations. The two methods showed at most a deviation of 30% from each other. Due to this high similarity and to save space only the results of Gillespie simulations are shown in figure 6.

ceRNA asymptotics

Here we assume that other than μ 's all the other parameters are equal between the two species. As a result the corresponding indices have been dropped in the following calculations. For the steady state values of concentrations we have:

$$\begin{aligned}
\bar{c}_i &= \frac{\mu_i \bar{m}_i \bar{s}}{\tau_{cR}^{-1}} \\
\bar{m}_i &= \frac{\alpha_{m_i}}{\tau_m^{-1} + q\phi\mu_i \bar{s}} \\
\alpha_s &= \tau_s^{-1} \bar{s} - (\beta + \gamma) \frac{\mu_1 \bar{m}_1 + \mu_2 \bar{m}_2}{\tau_{cR}^{-1}} \bar{s} + (\mu_1 \bar{m}_1 + \mu_2 \bar{m}_2) \bar{s}
\end{aligned}$$

which results in

$$\alpha_s = \tau_s^{-1} \bar{s} + q \left[\frac{\alpha_{m_1} \mu_1}{\tau_m^{-1} + q\phi\mu_1 \bar{s}} + \frac{\alpha_{m_2} \mu_2}{\tau_m^{-1} + q\phi\mu_2 \bar{s}} \right] \bar{s}$$

After some calculations we derive the following polynomial for sRNA concentration, \bar{s} :

$$\bar{s}^3 + (B_1 + B_2 + A_1 + A_2 - K)\bar{s}^2 + (B_1B_2 - K(B_1 + B_2) + A_1B_2 + A_2B_1)\bar{s} - KB_1B_2 = 0$$

with

$$K = \alpha_s \tau_s, \quad A_i = \frac{\tau_s \alpha_{m_i}}{\phi}, \quad B_i = \frac{\tau_s \lambda_i}{\phi}$$

furthermore in the following calculations $\lambda_i = \frac{1}{q\mu_i\tau_m\tau_s}$ and index T represents total sum over i , such that:

$$A_T = A_1 + A_2, \quad B_T = B_1 + B_2, \quad \lambda_T = \lambda_1 + \lambda_2, \quad \alpha_{m_T} = \alpha_{m_1} + \alpha_{m_2}$$

Expressing regime

In this regime $A_T \gg G \equiv \max(B_T, KB_T, B_1B_2)$ and we can simplify the polynomial equation by defining $\epsilon \equiv G/A_T$ and multiplying it by ϵ while keeping coefficients to first order:

$$\epsilon \bar{s}^3 + G\bar{s}^2 + GD\bar{s} - KB_1B_2\epsilon = 0$$

with $D \equiv \frac{A_1}{A_T}B_2 + \frac{A_2}{A_T}B_1$. Note that GD is of order $O(\epsilon^0)$ and does not require ϵ expansion. The equation for \bar{s} has the following asymptotic solutions:

$$\bar{s} = -\frac{G}{\epsilon}, -D, \frac{kB_1B_2}{GD}\epsilon$$

with the only positive solution being $\bar{s} = \frac{kB_1B_2}{GD}\epsilon = \frac{\alpha_s\tau_s}{\frac{\alpha_{m_1}}{\lambda_1} + \frac{\alpha_{m_2}}{\lambda_2}}$. In the limit of $\alpha_{m_2} = 0$, this reduces to the single species result in equation 27. Finally for mRNA we have $\bar{m} \sim \tau_m\alpha_{m_i}(1 - q\phi\tau_m\mu_i\bar{s}) \sim \tau_m\alpha_{m_i} + O(\epsilon)$ which is the expected result.

Repressing regime

In this regime $K \gg G \equiv \max(A_i, B_i)$ which is equivalent to $\phi\alpha_s \gg \alpha_{m_i}, \lambda_i$. Using this assumption we can simplify the polynomial equation by defining $\epsilon = G/K$ and multiplying it by ϵ while keeping coefficients to first order:

$$\epsilon \bar{s}^3 - (G - (A_T + B_T)\epsilon)\bar{s}^2 + ((B_1B_2 + A_1B_2 + A_2B_1)\epsilon - GB_T)\bar{s} - GB_1B_2 = 0$$

which has the following asymptotic solutions:

$$\bar{s} = -B_1, \quad -B_2, \quad \frac{G}{\epsilon} + x$$

with x being the solution for the following equation:

$$\epsilon\left(\frac{G}{\epsilon} + x\right)^3 - (G - (A_T + B_T)\epsilon)\left(\frac{G}{\epsilon} + x\right)^2 - GB_T\frac{G}{\epsilon} = 0$$

This results in $x = -A_T$. So the only positive solution is $\bar{s} \sim \frac{G}{\epsilon} - A_T = \frac{\tau_s}{\phi}(\phi\alpha_s - \alpha_{m_T})$ and $\bar{m}_i \sim \frac{\alpha_{m_i}\tau_m\lambda_i}{\phi\alpha_s - \alpha_{m_T}}$ which for $\alpha_{m_2} = 0$ reduce to our single species results in equation 32.

Crossover regime

Crossover regime is roughly where the repressing regime's line intersects with $\bar{s} = 0$, hence at this point we have $\alpha_{m_T} = \phi\alpha_s$

Acknowledgments

This work was supported by NIH Grants K25GM086909 (to P.M.), a Boston University Dean's Fellowship and a National Science Foundation Graduate Research Fellowship (NSF GRFP) under Grant No. DGE-0741448 (AHL).

References

1. Ambros V (2004) The functions of animal microRNAs. *Nature* 431: 350–5.
2. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. *Genes & development* 16: 1616–26.
3. Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75: 843–854.
4. Alvarez-Garcia I, Miska EA (2005) MicroRNA functions in animal development and human disease. *Development (Cambridge, England)* 132: 4653–62.
5. Miska EA (2005) How microRNAs control cell division, differentiation and death. *Current opinion in genetics & development* 15: 563–8.
6. Caldas C, Brenton JD (2005) Sizing up miRNAs as cancer genes. *Nature medicine* 11: 712–4.
7. Shenouda SK, Alahari SK (2009) MicroRNA function in cancer: oncogene or a tumor suppressor? *Cancer metastasis reviews* 28: 369–78.
8. Olde Loohuis NFM, Kos A, Martens GJM, Van Bokhoven H, Nadif Kasri N, et al. (2012) MicroRNA networks direct neuronal development and plasticity. *Cellular and molecular life sciences : CMLS* 69: 89–102.
9. Schratt G (2009) microRNAs at the synapse. *Nature reviews Neuroscience* 10: 842–9.
10. Fuchs Y, Steller H (2011) Programmed cell death in animal development and disease. *Cell* 147: 742–58.
11. Baehrecke EH (2003) miRNAs: Micro Managers of Programmed Cell Death. *Current Biology* 13: R473–R475.
12. Havelange V, Garzon R (2010) MicroRNAs: emerging key regulators of hematopoiesis. *American journal of hematology* 85: 935–42.
13. Chen CZ, Li L, Lodish HF, Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science (New York, NY)* 303: 83–6.
14. Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, et al. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* 118: 69–82.
15. Gottesman S, Storz G (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harbor perspectives in biology* 3.
16. Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. *Nature reviews Microbiology* 9: 578–89.
17. Bartel D (2004) MicroRNAs Genomics, Biogenesis, Mechanism, and Function. *Cell* 116: 281–297.

18. Pillai RS, Bhattacharyya SN, Filipowicz W (2007) Repression of protein synthesis by miRNAs: how many mechanisms? *Trends in cell biology* 17: 118–26.
19. Kai ZS, Pasquinelli AE (2010) MicroRNA assassins: factors that regulate the disappearance of miRNAs. *Nature structural & molecular biology* 17: 5–10.
20. Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS biology* 5: e229.
21. Levine E, Hwa T (2008) Small RNAs establish gene expression thresholds. *Current opinion in microbiology* 11: 574–9.
22. Mehta P, Goyal S, Wingreen NS (2008) A quantitative comparison of sRNA-based and protein-based gene regulation. *Molecular systems biology* 4: 221.
23. Mitarai N, Benjamin JAM, Krishna S, Semsey S, Csiszovszki Z, et al. (2009) Dynamic features of gene expression control by small regulatory RNAs. *Proceedings of the National Academy of Sciences of the United States of America* 106: 10655–9.
24. Jia Y, Liu W, Li A, Yang L, Zhan X (2009) Intrinsic noise in post-transcriptional gene regulation by small non-coding RNA. *Biophysical Chemistry* 143: 60–69.
25. Platini T, Jia T, Kulkarni RV (2011) Regulation by small RNAs via coupled degradation: Mean-field and variational approaches. *Physical Review E* 84.
26. Baker C, Jia T, Kulkarni R (2012) Stochastic modeling of regulation of gene expression by multiple small RNAs. *Physical Review E* 85.
27. Mukherji S, Ebert MS, Zheng GXY, Tsang JS, Sharp PA, et al. (2011) MicroRNAs can generate thresholds in target gene expression. *Nature genetics* 43: 854–9.
28. Levine E, McHale P, Levine H (2007) Small regulatory RNAs may sharpen spatial expression patterns. *PLoS computational biology* 3: e233.
29. Hao Y, Xu L, Shi H (2011) Theoretical analysis of catalytic-sRNA-mediated gene silencing. *Journal of Molecular Biology* 406: 195–204.
30. Gillespie DT (2007) Stochastic simulation of chemical kinetics. *Annual review of physical chemistry* 58: 35–55.
31. Kampen NV (2007) *Stochastic Processes in Physics and Chemistry, Third Edition* (North-Holland Personal Library). North Holland, 464 pp.
32. Swain PS (2004) Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *Journal of molecular biology* 344: 965–76.
33. Balázsi G, van Oudenaarden A, Collins JJ (2011) Cellular decision making and biological noise: from microbes to mammals. *Cell* 144: 910–25.
34. Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. *Cell* 123: 1025–36.
35. Kærn M, Elston T, Blake W, Collins J (2005) Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics* 6: 451–464.

36. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 99: 12795–800.
37. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146: 353–8.

Figure Legends

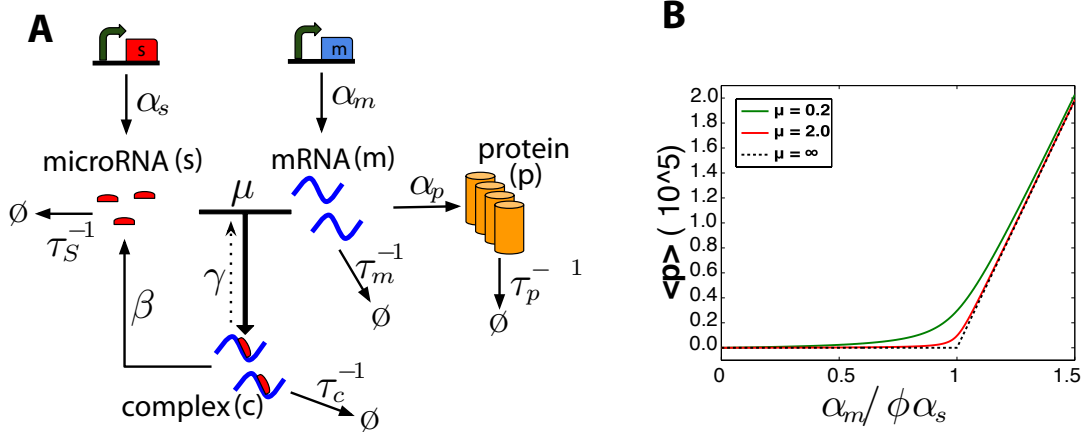


Figure 1. microRNA model: **A)** Schematic of the interactions between noncoding RNAs and mRNA. α_s and α_m are respectively the transcription rate of microRNA and mRNA. $\tau_m^{-1}, \tau_s^{-1}, \tau_c^{-1}, \tau_p^{-1}$ are respectively the degradation rates of mRNA, noncoding RNA, complex and protein. μ and γ are respectively the direct and reverse interaction rates between mRNA and noncoding RNA, and β is the catalyticity. **B)** Analytical results showing protein mean versus normalized transcription rate, $\frac{\alpha_m}{\phi \alpha_s}$, for different values of μ in the catalytic regime ($\beta = 10, \tau_c = 1$) where $\phi \equiv 1 + \beta \tau_c$. The dashed line is the theoretical result for infinitely large μ . The following parameters have been used in this plot: $\alpha_s = 3, \alpha_p = 4, \tau_s = 30, \tau_m = 10, \tau_p = 30, \gamma = 1$.

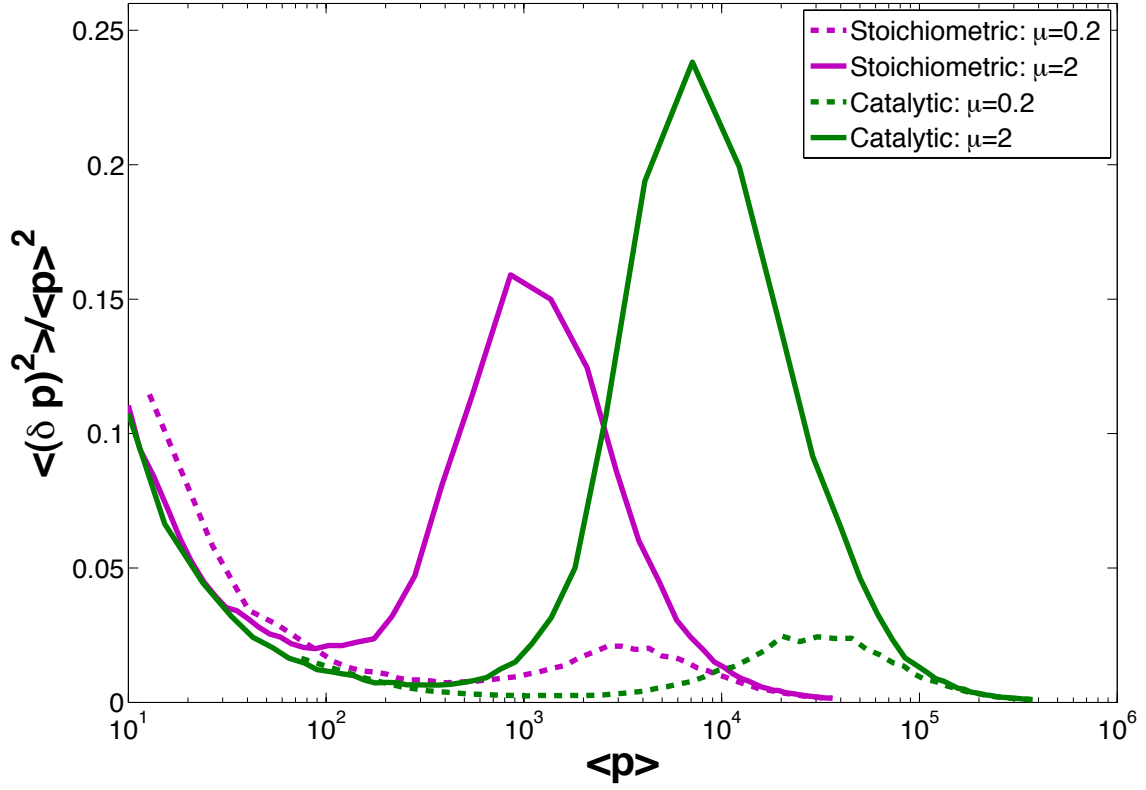


Figure 2. Protein noise in two regimes: Simulation results showing protein noise as a function of mean protein concentration for stoichiometric and catalytic interactions. For high μ (solid curves) there is higher noise in the crossover regime in the catalytic case. Same results are plotted for smaller μ (dashed curves) showing a decrease in noise at the crossover regime while maintaining the same overall behavior. Parameters same as in Figure 1B. Stoichiometric regime with $\beta = 0$ and catalytic regime with $\beta = 10$.

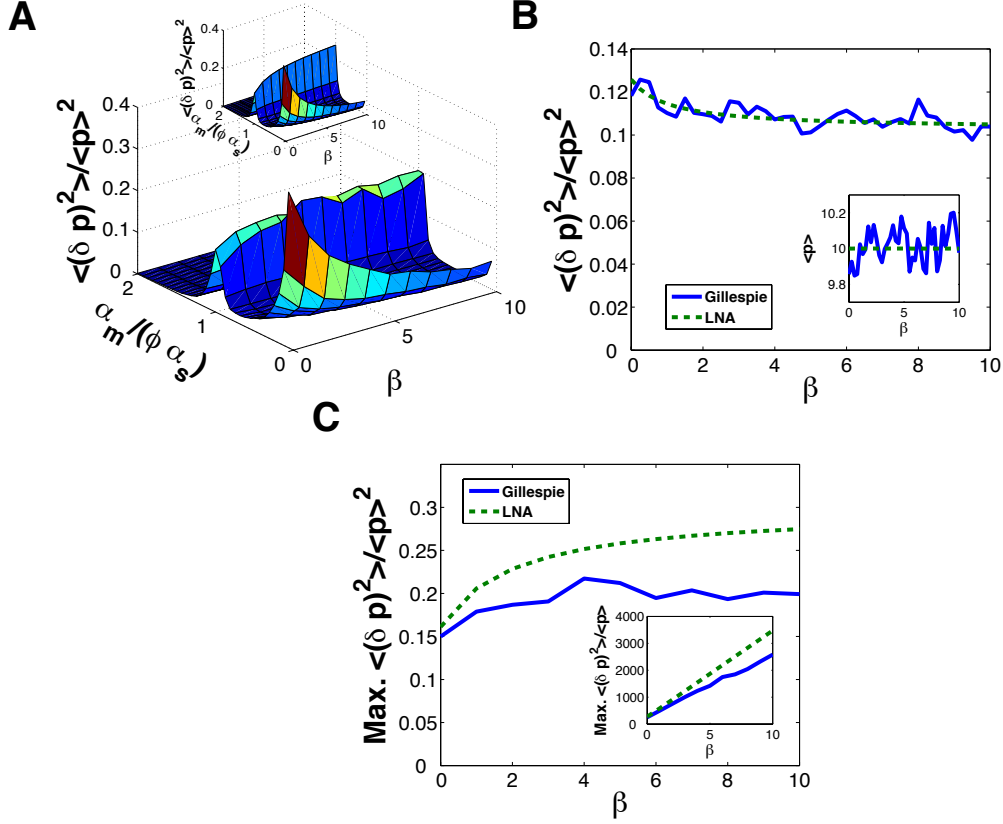


Figure 3. Effect of catalyticity on noise : **A)** Simulation results for noise as a function of β and $\frac{\alpha_m}{\phi \alpha_s}$. Inset: analytical results for the same system. Parameters same as Figure 1 with $\mu = 2$. **B)** Noise in the repressed regime as a function of β for constant protein mean ($\langle \bar{p} \rangle = 10$). Inset: protein mean as a function of β . **C)** Maximum of noise in the crossover regime as a function of β . Inset: maximum of Fano factor in the crossover regime as a function of β .

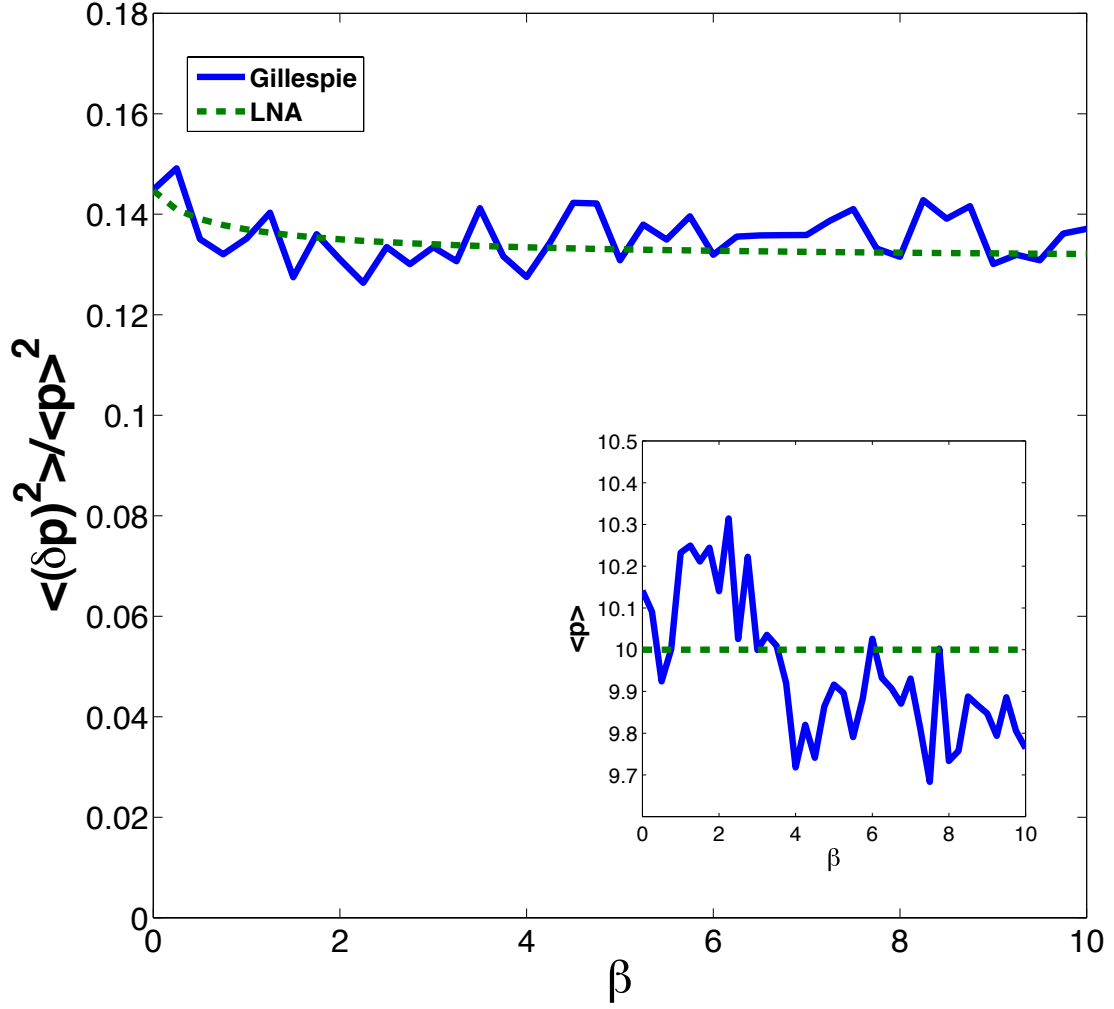


Figure 4. Catalyticity and Bursting: Noise in the repressed regime with bursting as a function of β for constant protein mean. For each data point α_m is chosen such that $\langle p \rangle = 10$. Furthermore $\alpha_m^{on} = 10, k_- = 1, k_+ = k_- (\frac{\alpha_m^{on}}{\alpha_m} - 1)$. The remaining parameters are same as in Figure 1 with $\mu = 2$. Inset: protein mean as a function of β .

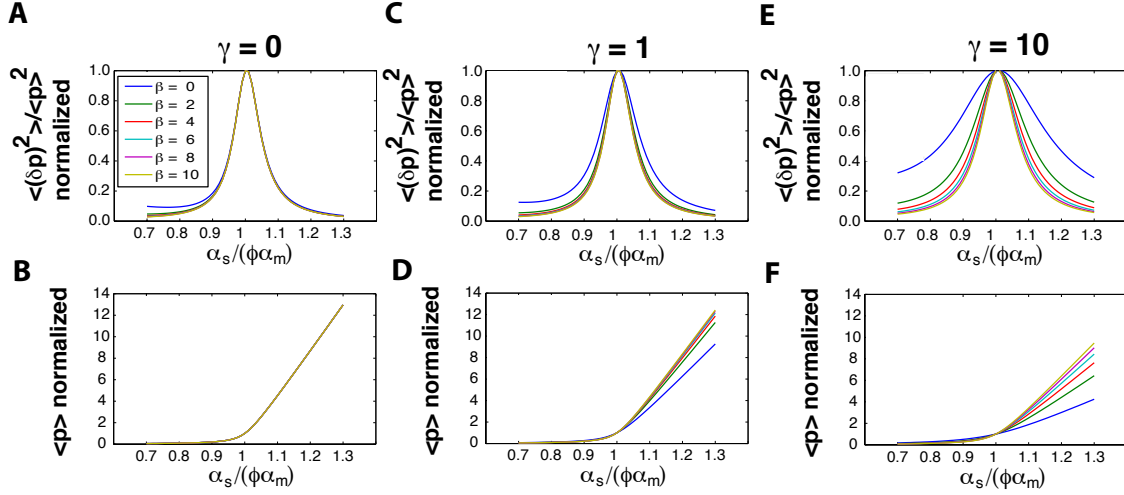


Figure 5. Scaling at crossover regime: Parameters same as Figure 1 with $\mu = 2$. **A)** Protein noise normalized to its value at $\alpha_m = \phi\alpha_s$ plotted as a function of $\frac{\alpha_m}{\phi\alpha_s}$ for $\gamma = 0$. Each line is a different value of β . Same legend for all figures. **B)** Protein mean normalized to its value at $\alpha_m = \phi\alpha_s$ plotted as a function of $\frac{\alpha_m}{\phi\alpha_s}$ for $\gamma = 0$. **C,D)** Graphs similar to **A,B** plotted for $\gamma = 1$. **E,F)** Graphs similar to **A,B** plotted for $\gamma = 10$.

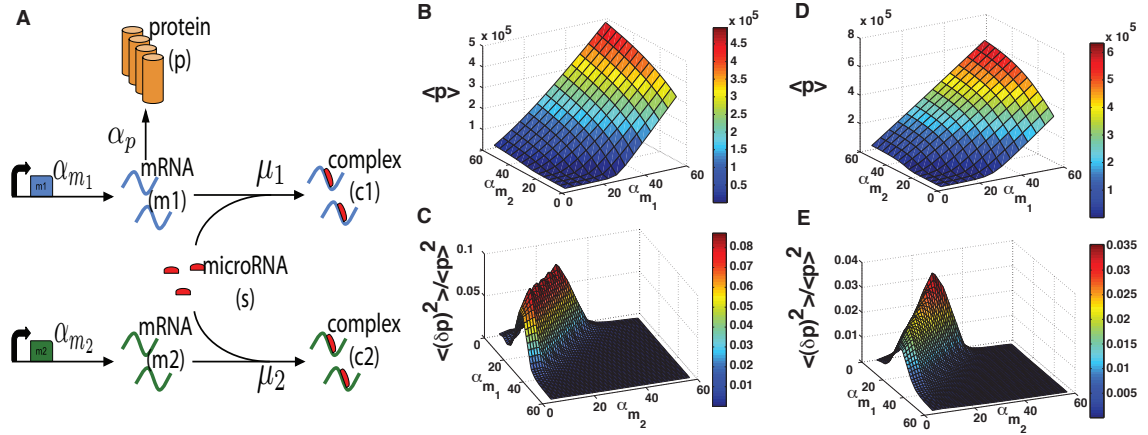


Figure 6. ceRNA hypothesis: **A)** Schematic of mRNA crosstalk through a shared pool of microRNAs. $\alpha_{m1,2}$ stand for transcription rate of each mRNA and $\mu_{1,2}$ correspond to interaction rates between mRNA and microRNA. The other interactions are as in Figure 1A. **B,C)** Protein mean (B) and noise (C) as a function of transcription rates of the two mRNAs with equal μ 's. Parameters same as in Figure 1 with $\mu_1 = \mu_2 = 2$. **D,E)** Protein mean (D) and noise (E) as a function of transcription rates of the two mRNAs with unequal μ 's. Parameters same as in Figure 1 with $\mu_1 = 0.2, \mu_2 = 2$. Noise surface plots have been smoothed and interpolated for better visibility.